2018

A cyclical social learning strategy for robust convention emergence

Yuchen Wang
University of Wollongong, yw808@uowmail.edu.au

Fenghui Ren
University of Wollongong, fren@uow.edu.au

Minjie Zhang
University of Wollongong, minjie@uow.edu.au

A cyclical social learning strategy for robust convention emergence

Abstract

Social conventions have been used as an efficient mechanism to facilitate coordination among agents. Establishing a convention in a decentralised manner has attracted much attention in the literature. Existing techniques on convention emergence are not robust. These techniques may establish sub-conventions under particular network structures. The emergence of sub-conventions indicates that agents in a society fail to conform to a single convention. As a result, the coordination among these agents is negatively affected. In this paper, we propose a strategy to avoid sub-conventions under diverse network structures. The proposed strategy requires agents to only have local views. We prove that a convention can be established using the proposed strategy. We also give empirical studies on the speed of convention emergence with various experimental settings.

Disciplines

Engineering | Science and Technology Studies

Publication Details

Wang, Y., Ren, F. & Zhang, M. (2018). A cyclical social learning strategy for robust convention emergence. Proceedings - 2018 IEEE International Conference on Agents (ICA) (pp. 120-125). United States: IEEE.

A Cyclical Social Learning Strategy for Robust Convention Emergence

Yuchen Wang, Fenghui Ren, and Minjie Zhang
School of Computing and Information Technology
University of Wollongong
Wollongong, Australia
yw808@uowmail.edu.au, fren@uow.edu.au, minjie@uow.edu.au

***Abstract*—Social conventions have been used as an efficient mechanism to facilitate coordination among agents. Establishing a convention in a decentralised manner has attracted much attention in the literature. Existing techniques on convention emergence are not robust. These techniques may establish sub-conventions under particular network structures. The emergence of sub-conventions indicates that agents in a society fail to conform to a single convention. As a result, the coordination among these agents is negatively affected. In this paper, we propose a strategy to avoid sub-conventions under diverse network structures. The proposed strategy requires agents to only have local views. We prove that a convention can be established using the proposed strategy. We also give empirical studies on the speed of convention emergence with various experimental settings.**

***Index Terms*—Social conventions, sub-conventions, multi-agent learning**

I. Introduction

Social conventions, such as driving rules, play an important role in regulating individual behaviours towards a global consensus in human societies [1]. The concept of social conventions has been widely used in the literature of multi-agent systems to improve coordination among agents. A convention restricts agents' behaviour to one particular action [2]. Conformity to a convention can reduce social conflicts and sustain social order in an agent society.

To establish a convention in an agent society, there are two kinds of approaches [3]. The first one is the top-down approach which requires a central authority to specify how agents should behave. The second one is the bottom-up approach where there is no central authority. Using the bottom-up approach, a convention emerges as a natural result of agents' repeated local interactions. Compared with the top-down approach, the bottom-up approach can offer a wider range of applications as an agent society is usually organised in a decentralised manner.

A number of bottom-up-based techniques have been proposed to facilitate convention emergence through agents' local learning actions under diverse network structures [4]–[6]. Learning has been widely adopted as a basic decision-making mechanism for agents to choose their actions [5]. Network structures are fundamental in the process of convention emergence [3]. A network structure determines which pair of agents can interact with each other. However, using the existing techniques, sub-conventions may emerge in an agent society under particular network structures [5], [7], [8]. The emergence of sub-conventions indicates that different conventions emerge in different regions of an agent society. In other words, the entire population fails to conform to a single convention, which might cause social conflicts and disturb social order.

In this paper, we focus on the problem of avoiding sub-conventions under diverse network structures. Identifying reasons for the emergence of sub-conventions is important in developing techniques to avoid sub-conventions. Researchers have identified one reason for the emergence of sub-conventions. The reason is that under particular network structures, the agents in one region interact with each other more often than with agents outside this region [5], [9]. As a result, in this region, agents are likely to reinforce a strategy different from the strategies reinforced in other regions. Based on the above reason, in [9], the authors claim that the problem of avoiding sub-conventions has to be solved through topological reconfigurations. This solution requires an agent to have abilities to change the agent's position, and learn from agents beyond the agent's neighbours. However, commonly, an agent has only local views, i.e., the agent cannot change its position, and can only learn from its neighbours. Therefore, how to avoid sub-conventions under diverse network structures with agents' local views remains a challenging problem.

To tackle this problem, we identify another reason for the emergence of sub-conventions. When sub-conventions emerge, an agent who connects two sub-convention regions will find that some of the agent's neighbours have learned strategies different from the agent. The agent's strategy is hence unacceptable because a consensus among the agent's neighbours (refer to as neighbouring consensus hereafter) is not yet reached. However, using current techniques [4]–[6], the agent does not check whether a neighbouring consensus is reached. As a result, the sub-conventions remain stable. In the real world, when a group of people are learning to reach a consensus, after finishing learning, a person would check whether a neighbouring consensus is reached. If the neighbouring consensus is not reached, the person would start to learn again with the neighbours whose strategies are different from the person. The learning and checking of neighbouring consensus happen cyclically until each person finds that the neighbouring consensus of him/her is reached. Based on the above intuition, we propose a Cyclical Social Learning (CSL) strategy which requires agents to only have local views. In CSL, an agent cyclically performs learning and checking of neighbouring consensus. We prove that using CSL, the neighbouring consensus of all agents can be reached. As a result, sub-conventions are avoided and a single convention is established. In addition to the theoretical work, we also conduct experiments to verify the effectiveness of CSL, and empirically study the speed of convention emergence.

The rest of this paper is organised as follows. Section 2 formulates the problem of avoiding sub-conventions. Section 3 introduces the proposed cyclical social learning strategy. Section 4 shows experimental studies on the proposed strategy. Section 5 reviews the related work. Finally, Section 6 concludes this paper.

II. Problem Formulation

This section formulates the problem we consider in this paper. We first introduce the concept of pure coordination games and social conventions. Then, we describe the emergence of sub-conventions in a networked agent society.

A. Pure Coordination Games and Social Conventions

We follow the game-theoretic framework [2] to study the emergence of social conventions and sub-conventions. In this framework, agents in a society strive to establish a convention through repeated local learning interactions. The learning interactions between agents are framed as 2-player m-action pure coordination games. A convention is said to be established when all agents have learned to adopt the same action. By contrast, when the agents have learned to adopt different actions, sub-conventions are said to be established. In a 2-player m-action pure coordination game, there are m number of equally good Nash equilibria, which indicates m possible conventions. Table I shows the payoff matrix of a typical pure coordination game where $m = 2$. This game has two equally good equilibria: both agents choose the action a_1 or both agents choose the action a_2.

TABLE I
Payoff matrix of a 2-player 2-action pure coordination game

	Action a_1	Action a_2
Action a_1	1, 1	-1, -1
Action a_2	-1,-1	1, 1

B. Networked Agent Societies and Sub-Conventions

A networked agent society can be represented as a graph $G = (V, E)$ where $V = \{v_1, ..., v_n\}$ is a set of vertices and each vertex represents an agent, $E \subseteq V \times V$ is a set of edges and each edge connects two agents. The neighbours of an agent i, denoted as $N(i)$, compose a set of agents where $N(i) = \{v_j | (v_i, v_j) \in E\}$. Commonly, in a networked agent society, an agent can only interact with its neighbours.

Researchers have identified that the underlying network structure of a networked agent society plays a central role in the emergence of sub-conventions [5], [7]. Figure 1 shows an example of the emergence of sub-conventions under a fully-connected-star network. In this example, the agents A, B, C and D have converged to an action strategy (represented by the nodes with light colour), while the agents E, F, G and H have converged to another action strategy (represented by the nodes with dark colour). This convergent result indicates the emergence of two sub-conventions.

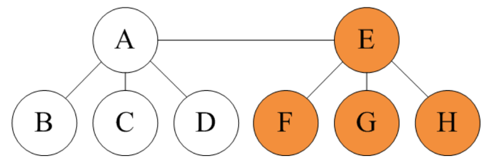


Fig. 1. The emergence of sub-conventions in a fully-connected-star network

In this paper, we study how to avoid sub-conventions under diverse network structures. Avoiding sub-conventions means establishing a single convention in an agent society. To measure the degree of the establishment of a convention, we compute the frequency of the action adopted by the majority of agents: $F = \frac{\max_a c(a)}{n}$ where n is the total number of agents, $c(a)$ is the number of agents that adopt the action a. A convention is established when $F = 1$.

III. Establishing a Convention while Avoiding Sub-conventions

In this section, we introduce the proposed Cyclical Social Learning (CSL) strategy. First, we give the interaction process of an agent. Then, we provide a theorem which shows that agents using CSL can avoid sub-conventions. Lastly, we discuss the speed of convention emergence using CSL.

A. Cyclical Social Learning Strategy

We consider a networked agent society. Each agent in the society interacts repeatedly and simultaneously with the agent's neighbours. The interaction process of an agent is presented in Algorithm 1. Each agent contains a learning state whose possible values are "Learning" and "Learned". Before the interaction process begins, an agent first initialises its learning state to "Learning" (Line 1). Then, the agent starts to interact with its neighbours. In each interaction, depending on the value of the learning state, the agent performs one of the two procedures described below:

1) Procedure in "Learning" state: When the agent is in "Learning" state, the agent first selects an action based on the agent's current learning information using a specific learning algorithm (Line 4). Then, the agent plays the selected action with neighbours who are also in "Learning" state, and updates the learning information using the received reward (Lines 5-8). The play between agents is framed as a pure coordination game (refer to Section II-A). After the agent has learned with a fixed number of episodes, the agent has accumulated certain amounts of learning information. The learning state is then changed to "Learned" (Lines 9-10). The fixed number of learning episodes could be set to an empirical value.

Algorithm 1: The interaction process of an agent i

1 Initialises the learning state ls_i to "*Learning*";
2 **for** *each episode* **do**
3   **if** ls_i = *"Learning"* **then**
4   Selects an action a_i using a learning algorihtm with ϵ-greedy exploration;
5   **for** *each neighbour* $j \in N(i)$ **do**
6   Observes the learning state ls_j of the neighbour j;
7   **if** ls_j = *"Learning"* **then**
8   Plays the action a_i with j and updates the learning information using the received reward;
9   **if** *has learned with a fixed number of episodes* **then**
10   Sets the learning state ls_i to "*Learned*";
11   **else if** ls_i = *"Learned"* **then**
12   Obtains the strategy a_i^* based on the accumulated learning information;
13   **for** *each neighbour* $j \in N(i)$ **do**
14   Observes the learning state ls_j of the neighbour j;
15   **if** ls_j = *"Learned"* **then**
16   Observes the strategy a_j^* of the neighbour j;
17   **if** $a_i^* \neq a_j^*$ **then**
18   Resets the learning information;
19   Sets the learning state ls_i to "*Learning*";

2) Procedure in "Learned" state: When the agent is in "Learned" state, the agent first obtains its current strategy based on its learning information (Line 12). Next, for each neighbour who is also in "Learned" state, the agent compares its strategy with the neighbour's strategy. If a different strategy is observed, the agent resets its learning information, and sets its learning state to "Learning" (Lines 17-19).

B. Avoiding Sub-Conventions in Networked Agent Societies

We consider an agent society where sub-conventions have emerged. To show how agents using CSL avoid sub-conventions, let $P(F^{(l)} < 1)$ denote the possibility of being in a sub-convention situation after the agents perform the lth learning procedure. Performing the learning procedure for t consecutive times can be regarded as an operator $L^{(t)}$. $L^{(t)}$ is applied to the agent society to update $P(F^{(l)} < 1)$. Then, we need to prove that there exists a number t such that $L^{(t)}$ is a contraction to $P(F^{(l)} < 1)$. In this condition, the possibility of being in a sub-convention situation will decrease to zero when $L^{(t)}$ is repeatedly applied to the agent society. To do so, we first give the definitions and lemma below. The term *sub-convention region* $\overline{a_i}$ is used to indicate a region of agents who adopt the action a_i.

Definition 1: (*Bridging Agent*). In an agent society where sub-conventions have emerged, the bridging agents of a sub-convention region $\overline{a_i}$, denoted as $b(\overline{a_i})$, are composed of two parts of agents: α and β. The agents α are in $\overline{a_i}$, and connect to other sub-convention regions. The agents β are in other sub-convention regions and connect to α.

Definition 2: (*Expansion Range*). In an agent society where sub-conventions have emerged, the expansion range of a sub-convention region $\overline{a_i}$, denoted as $er(\overline{a_i})$, is the distance between $\overline{a_i}$ and the agent furthest from $\overline{a_i}$.

Consider a sub-convention region $\overline{a_i}$, the bridging agents will observe that their strategies are different from some of their neighbours' strategies during the procedure in "Learned" state. Then, these bridging agents reset their learning information and start to perform the next learning procedure. In this learning procedure, these bridging agents have a probability of converging to a_i. The convergence to a_i can be regarded as an expansion of the sub-convention region $\overline{a_i}$. To show how agents can avoid sub-conventions, we are going to prove that the sub-convention region $\overline{a_i}$ has a probability of expanding to all the agent society.

Lemma 1: In an agent society where sub-conventions have emerged, after the agents perform the learning procedures for $er(\overline{a_i})$ times, the agent society will converge to a_i with a minimum probability $P(F^{(er(\overline{a_i}))} = 1)$.

Proof: Let $P(b(\overline{a_i}) \to a_i)$ be the probability that $b(\overline{a_i})$ converge to the strategy a_i after one learning procedure, m be the number of actions, $c(b(\overline{a_i}))$ be the number of $b(\overline{a_i})$. The probability that all $b(\overline{a_i})$ randomly choose a_i as their initial actions is $(1/m)^{c(b(\overline{a_i}))}$. This probability is also the minimum value of $P(b(\overline{a_i}) \to a_i)$ because when all $b(\overline{a_i})$ choose a_i as their initial actions, a_i will be reinforced during learning and $b(\overline{a_i})$ will converge to a_i[1]. Hence, after performing the learning procedures for $er(\overline{a_i})$ times, a_i will expand to all agents with a probability $P(F^{(er(a_i))} = 1) = \prod_{j=1}^{er(\overline{a_i})}[P(b_j(\overline{a_i}) \to a_i]$ where $b_j(\overline{a_i})$ denotes the bridging agents of the sub-convention region $\overline{a_i}$ after performing the jth learning procedure. $P(b_0(\overline{a_i}) \to a_i)$ is set to 1 to indicate the already emerged sub-convention situation. ■

Theorem 1: In an agent society where sub-conventions have emerged, there exists a number $t = \max_{a_i} er(\overline{a_i})$, such that $L^{(t)}$ is a contraction to $P(F^{(l)} < 1)$.

Proof: As $t = \max_{a_i} er(\overline{a_i})$, based on Lemma 1, after agents perform learning procedures for t consecutive times (i.e., applying $L^{(t)}$ once), all agents can converge to one strategy with a minimum probability $P(F^{(l+t)} = 1) = \sum_{i=1}^{m}\prod_{j=1}^{t} P(b_j(\overline{a_i}) \to a_i)$. Hence, $P(F^{(l+t)} = 1) > 0$ and the following inequality holds:

$$L^{(t)}P(F^{(l)} < 1) \leq \gamma P(F^{(l)} < 1) \tag{1}$$

[1]The effect of exploration is cancelled out by that the agents have the same possibility of deviating from a_i to other actions and from other actions to a_i.

where $\gamma = P(F^{(l+t)} < 1) = 1 - P(F^{(l+t)} = 1)$ and $\gamma < 1$. This means that $L^{(t)}$ is a contraction to the probability $P(F^{(l)} < 1)$ by the factor of γ.

■

C. Discussion on the Speed of Convention Emergence

Based on the above section, the speed of convention emergence depends on how many learning procedures required to make a sub-convention region expand to the entire agent society. For a sub-convention region $\overline{a_i}$, let $\mathring{a}$ be the agent furthest from $\overline{a_i}$, we consider two factors that would influence the speed of expanding $\overline{a_i}$ to $\mathring{a}$:

1) The diameter of a network: The diameter of a network indicates the shortest distance between the two most distant nodes in the network. We expect that when the diameter of a networked agent society is shorter than societies with longer diameter, the $er(\overline{a_i})$ of the sub-convention region $\overline{a_i}$ would also be shorter. As a result, the minimum number of learning procedures required of expanding $\overline{a_i}$ to $\mathring{a}$ is smaller, and the speed of convention emergence would be faster.

2) The average degree of a network: In a networked agent society, the degree of an agent is the number of its neighbours. The average degree of a network is the average degree of all agents. We expect that when the average degree of a networked agent society is higher than societies with lower average degree, there would be more paths of expanding $\overline{a_i}$ to $\mathring{a}$ (a path is a sequence of agents involved during an expansion). As a result, the probability of expanding $\overline{a_i}$ to $\mathring{a}$ is bigger, and the speed of convention emergence would be faster.

IV. Experimental Results

In this section, we evaluate the proposed CSL strategy with various experimental settings.

A. Experimental Settings

1) Network Structure Settings: Previous work has identified that sub-conventions may emerge under these four kinds of networks: (1) scale-free networks [7], (2) fully-connected-star networks [7], (3) community networks [8] and (4) isolated agent groups with infrequent interactions between agents of different groups [5]. Under each kind of network, the emergence of sub-conventions is influenced by an important topological parameter. The influential parameters and their values are shown in Table II. We generate networked agent societies under each kind of network with a particular influential parameter. The values of other topological parameters under each network are either fixed or randomly selected.

2) Investigated Issues: We investigate three issues to evaluate the proposed CSL strategy. In the investigation, unless specifically stated, Q-Learning [10] is used as the default learning algorithm due to its popularity. The number of actions is set to 2, indicating 2 different ways to establish a convention.

Issue 1: Success ratio. The most fundamental issue is the success ratio of establishing a convention (and hence avoiding sub-conventions) using CSL. For comparison, we also apply Social Learning (SL) [5], which is a representative and commonly used technique for convention emergence. This issue is investigated to experimentally verify Theorem 1.

TABLE II
INFLUENTIAL TOPOLOGICAL PARAMETERS FOR EACH NETWORK

Networks	Parameters	Values	Meanings
Scale-free networks	n	{20, 40, 60, 80}	Number of agents
Fully-connected-star networks	c	{2, 4, 6, 8}	Number of clusters. Each cluster is a star network.
Community networks	σ	{0.6, 0.7, 0.8, 0.9}	Separation degree, which denotes relative ratio of intra-community neighbours to total number of neighbours of each agent.
Isolated agent groups	p	{0.01, 0.05, 0.1, 0.2}	Probability of interaction between agents of different groups

Issue 2: Speed of convention emergence. We are also interested in the speed of convention emergence using CSL. The number of learning procedures required to establish a convention is recorded. The diameter and average degree of a network are also recorded to investigate the influence of these factors on the speed of convention emergence.

Issue 3: Number of bridging agents and their converging results. The number of bridging agents indicates the number of agents who observe different strategies among neighbours after one learning procedure. In the next learning procedure, these agents will either converge to the same strategy or fail to do so. Looking into these statistics would also give insights into the proposed CSL strategy. We record the values of these variables when sub-conventions emerge.

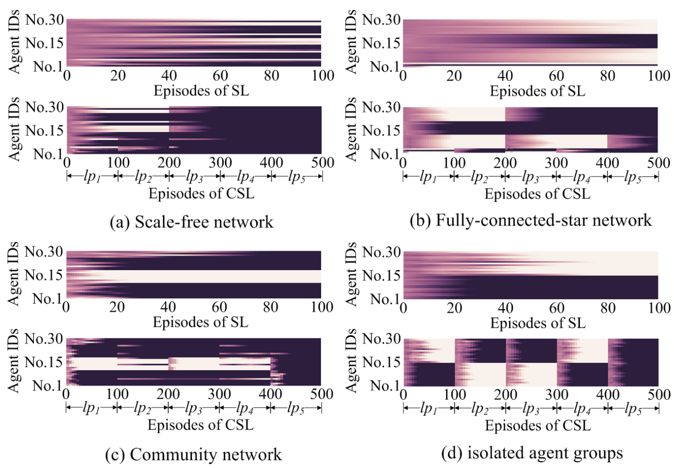
TABLE III
SUCCESS RATIO OF ESTABLISHING A CONVENTION UNDER FOUR KINDS OF NETWORKS (AVERAGE OVER 1,000 RUNS)

Scale-free networks			Fully-connected-star networks		
n	SL	CSL	c	SL	CSL
20	32.9%	100%	2	52.4%	100%
40	18.5%	100%	4	34.9%	100%
60	4.1%	100%	6	22.2%	100%
80	1.6%	100%	8	13.3%	100%
Community networks			**Isolated agent groups**		
σ	SL	CSL	p	SL	CSL
0.6	88.1%	100%	0.2	85.2%	100%
0.7	49.8%	100%	0.1	68.4%	100%
0.8	23.7%	100%	0.05	56.7%	100%
0.9	19.2%	100%	0.01	53.2%	100%

B. Results and Analysis

1) Avoiding Sub-conventions: Table III shows the success ratio of establishing a convention in agent societies under the four kinds of networks. We can see that when agents use SL, the success ratio is smaller than 100% under all networks. These results indicate the emergence of sub-conventions, and failure of establishing a convention. Also, the success ratio of SL is significantly influenced by the influential topological parameters under each kind of network. In particular, the success ratio decreases in situations when n increases in scale-free networks, when c increases in fully-connected-star networks, when σ increases in community networks, and when

p decreases in isolated agent groups. As introduced in Section II-B, the sub-conventions might be caused by some particular network structures. We can expect that these particular structures are more likely to cause sub-conventions when the network becomes bigger or more separate. In contrast, when agents use CSL, the success ratio reaches 100% under all settings. The values of the influential parameters have no impact on the success ratio. These results experimentally verify the proposed Theorem 1. In summary, CSL is robust for avoiding sub-conventions and establishing a single convention under diverse network structures.



(a) Scale-free network (b) Fully-connected-star network

(c) Community network (d) isolated agent groups

Fig. 2. Examples of dynamics of the probability to play the action a_1 under four kinds of networks. For each network, the dynamics of SL and CSL are shown on above and below respectively. lp_i denotes the ith learning procedure of CSL.

To better understand how sub-conventions can be avoided using CSL, we present typical examples of the learning dynamics in a population of 30 agents under each of the four kinds of networks. In these examples, Win-or-Learn-Fast with Policy Hill-Climbing (WoLF-PHC) [11] is used as the learning algorithm due to the ease of demonstrating the agents' policies. To use CSL, we set the number of episodes in each learning procedure to 100. Figures 2(a)-2(d) show the learning dynamics of the probability to play the action a_1 under each kind of network. In each dynamics, an agent is represented by a square. The x-axis denotes the number of episodes, and the y-axis denotes the policy of each agent in one episode. Darker colour indicates a larger probability to play the action a_1. When SL is applied, we can see that under each kind of network, the colour of the agents gradually becomes different (darker or lighter). This indicates that the whole population fails to establish a convention, and sub-conventions emerge in different regions of an agent society. We only present the results within the first 100 episodes. It could be noticed that the convergence degree does remain stable in all subsequent episodes. By contrast, when CSL is applied, we can see that after each learning procedure, some agents reset their learning information and start to learn again in the next learning procedure. These agents do this because they find that some of their neighbours adopt different strategies, and hence another learning procedure is initiated. As the agents cyclically perform learning procedures, all agents gradually converge to the same policy. As a result, sub-conventions are avoided and a convention is established.

TABLE IV
THE DIAMETER, AVERAGE DEGREE AND LEARNING PROCEDURES SPENT BEFORE A CONVENTION IS ESTABLISHED UNDER FOUR KINDS OF NETWORKS (AVERAGE OVER 1,000 RUNS)

Scale-free networks				Fully-connected-star networks			
n	*dia*	*deg*	Speed	*c*	*dia*	*deg*	Speed
20	5.89	1.91	6.60	2	3	1.9	3.38
40	8.24	1.95	23.89	4	4	2	8.12
60	9.14	1.96	32.32	6	5	2	14.42
80	9.7	1.97	115.04	8	6	2	42.54
Community networks				Isolated agent groups			
σ	*dia*	*deg*	Speed	*p*	*dia*	*deg*	Speed
0.6	2	40.57	1.13	0.2	2	123	1.17
0.7	2	35.27	2.16	0.1	2	110	1.50
0.8	3	31.98	4.51	0.05	2	104	1.75
0.9	3	28.92	7.76	0.01	2	100	1.89

2) Speed of Convention Emergence: We investigate the speed of convention emergence using CSL by measuring the number of learning procedures spent before a convention emerges. As discussed in Section III-C, we expect that the diameter and average degree of a network would influence the speed. Table IV shows the diameter, average degree, and average learning procedures spent under each kind of network. We can see that under scale-free networks and fully-connected-star networks, the average degree is similar with different influential parameters. The speed would be mainly influenced by the diameter. Under both kinds of networks, shorter diameter results in less learning procedures spent before a convention is established. This verifies our expectation that when the diameter of an agent society is shorter, less learning procedures are required to expand a sub-convention region to the entire population. As a result, the speed of convention emergence is faster. Under community networks and isolated agent groups, the diameter is similar with different influential parameters. The speed would be mainly influenced by the average degree. Under both kinds of networks, larger average degree results in less learning procedures spent. This also verifies our expectation that when the average degree is larger, there would be more paths of expanding a sub-convention region to the entire population, which accelerates the speed of convention emergence.

3) Number of Bridging Agents and Their Converging Results: Table V shows the average number of bridging agents (Column *c(b)*) and their converging results. The columns $P(b \rightarrow a_1)$ and $P(b \rightarrow a_2)$ indicate the probability of converging to the action strategy a_1 and a_2 respectively. We can see that under scale-free networks and fully-connected-star networks, the probability of converging to the same action strategy among the bridging agents is high ($P(b \rightarrow a_1) + P(b \rightarrow a_2) > 90\%$). This means that a consensus is likely to be reached among the bridging agents. Under community networks and isolated agent groups, the convergence probability decreases when the communities become more separate and when the interaction probability decreases respectively. In all settings, the bridging agents converge to each strategy with almost equal probabilities. This makes sense because the pure coordination game does not provide preference regarding which strategy should be the convention.

TABLE V
NUMBER OF BRIDGING AGENTS AND THEIR CONVERGING RESULTS (AVERAGE OVER 10,000 SUB-CONVENTION SITUATIONS)

Scale-free networks				Fully-connected-star networks			
n	$c(b)$	$P(b\rightarrow a_1)$	$P(b\rightarrow a_2)$	c	$c(b)$	$P(b\rightarrow a_1)$	$P(b\rightarrow a_2)$
20	3.27	0.502	0.497	2	6.0	0.498	0.502
40	4.62	0.498	0.499	4	10.04	0.455	0.429
60	4.15	0.499	0.493	6	9.61	0.452	0.465
80	4.10	0.483	0.485	8	9.27	0.471	0.457
Community networks				Isolated agent groups			
σ	$c(b)$	$P(b\rightarrow a_1)$	$P(b\rightarrow a_2)$	p	$c(b)$	$P(b\rightarrow a_1)$	$P(b\rightarrow a_2)$
0.6	79.97	0.431	0.437	0.2	200	0.399	0.393
0.7	73.51	0.257	0.262	0.1	200	0.317	0.325
0.8	65.54	0.183	0.184	0.05	200	0.285	0.287
0.9	52.14	0.153	0.153	0.01	200	0.258	0.261

V. RELATED WORK

Many techniques had been proposed to facilitate the emergence of a convention in an agent society. Shoham and Tennenholtz [2] proposed the Highest Cumulative Reward (HCR) rule to study convention emergence. An agent using HCR chose the action strategy which yielded the highest reward in past m iterations. Sen and Airiau [5] proposed a social learning framework to facilitate convention emergence through repeated learning interactions. The above work, however, studied convention emergence under unstructured agent societies where each agent could interact with all other agents.

In the real world, an agent society is usually organised under a structured network. Delgado [3] studied convention emergence under complex networks. This work proposed the Generalised Simple Majority rule which provided analytical evidence of convergence to a convention. Yu *et al.* [4] proposed a collective learning framework to study the impact of agents' collective learning behaviours on convention emergence under various network structures. Vouros [12] studied convention emergence in a setting which required agents to accomplish multiple tasks simultaneously with operational constraints. However, the main results of the above work were conducted under restricted kinds of networks (e.g., small-world networks and scale-free networks). Also, the above work measured the emergence of a convention using the 90% convergence criterion proposed by Kittock [13]. Using this criterion, in an agent society, a convention was said to have emerged if at least 90% of the agents in this society had converged to the same strategy. However, as argued in [9], the 90% convergence should not be considered as an appropriate criterion to measure the emergence of a convention (at least a robust one). By the definitions of conventions proposed in [1], [2], a convention should be shared by 100% of the agents in an agent society. When the agents failed to conform to a single convention, sub-conventions were said to have emerged.

Researchers had studied the emergence of sub-conventions under a variety of network structures. Sen and Airiau [5] identified the phenomenon of sub-conventions in isolated agent groups where agents in different groups interacted only infrequently. Villatoro *et al.* [7] identified the phenomenon of sub-conventions under scale-free networks and fully-connected-star networks. Airiau *et al.* [5] provided some explanations of how sub-conventions emerged in scale-free networks. Hu and Leung [8] showed that how sub-conventions could be utilised under community networks to coordinate agents' actions in each community. However, the above studies did not provide solutions to avoid sub-conventions. Villatoro [9] proposed a solution to avoid sub-conventions. This solution required agents to have global views. However, in an agent society, an agent usually had only local views, which made Villatoro's solution inapplicable in many circumstances.

VI. CONCLUSION

In this paper, we propose a Cyclical Social Learning (CSL) strategy for robust convention emergence. We provide a theorem which shows that using CSL, a convention is guaranteed to be established in diverse networked agent societies. The speed of convention emergence using CSL is also empirically studied. In future, we plan to analytically study the speed of convention emergence using CSL.

ACKNOWLEDGEMENT

This research is supported by a DECRA Project (DP140100007) from Australia Research Council (ARC), a UPA and an IPTA scholarships from University of Wollongong, Australia.

REFERENCES

[1] H. P. Young, "The economics of convention," *The Journal of Economic Perspectives*, vol. 10, no. 2, pp. 105–122, 1996.

[2] Y. Shoham and M. Tennenholtz, "On the emergence of social conventions: modeling, analysis, and simulations," *Artificial Intelligence*, vol. 94, no. 1, pp. 139–166, 1997.

[3] J. Delgado, "Emergence of social conventions in complex networks," *Artificial Intelligence*, vol. 141, no. 1-2, pp. 171–185, 2002.

[4] C. Yu, M. Zhang, F. Ren, and X. Luo, "Emergence of social norms through collective learning in networked agent societies," in *Proceedings of the 12th International Conference on Autonomous Agents and Multi-Agent Systems*, 2013, pp. 475–482.

[5] S. Airiau, S. Sen, and D. Villatoro, "Emergence of conventions through social learning," *Autonomous Agents and Multi-Agent Systems*, vol. 28, no. 5, pp. 779–804, 2014.

[6] J. Hao, J. Sun, G. Chen, Z. Wang, C. Yu, and Z. Ming, "Efficient and robust emergence of norms through heuristic collective learning," *ACM Transactions on Autonomous and Adaptive Systems (TAAS)*, vol. 12, no. 4, p. 23, 2017.

[7] D. Villatoro, S. Sen, and J. Sabater-Mir, "Topology and memory effect on convention emergence," in *Proceedings of the 2009 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology-Volume 02*, 2009, pp. 233–240.

[8] S. Hu and H.-f. Leung, "Achieving coordination in multi-agent systems by stable local conventions under community networks," in *Proceedings of the 26th International Joint Conferences on Artificial Intelligence*, 2017, pp. 4731–4737.

[9] D. Villatoro, J. Sabater-Mir, and S. Sen, "Social instruments for robust convention emergence," in *Proceedings of the 22nd International Joint Conferences on Artificial Intelligence*, vol. 11, 2011, pp. 420–425.

[10] C. J. Watkins and P. Dayan, "Q-learning," *Machine Learning*, vol. 8, no. 3-4, pp. 279–292, 1992.

[11] M. Bowling and M. Veloso, "Multiagent learning using a variable learning rate," *Artificial Intelligence*, vol. 136, no. 2, pp. 215–250, 2002.
[12] G. A. Vouros, "Learning conventions via social reinforcement learning in complex and open settings," in *Proceedings of the 16th Conference on Autonomous Agents and Multi-Agent Systems*, 2017, pp. 455–463.
[13] J. E. Kittock, "Emergent conventions and the structure of multi-agent systems," in *Proceedings of the 1993 Santa Fe Institute Complex Systems Summer School*, vol. 6, 1993, pp. 1–14.